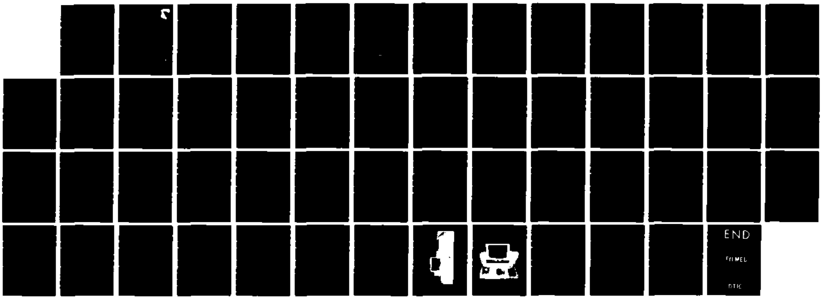
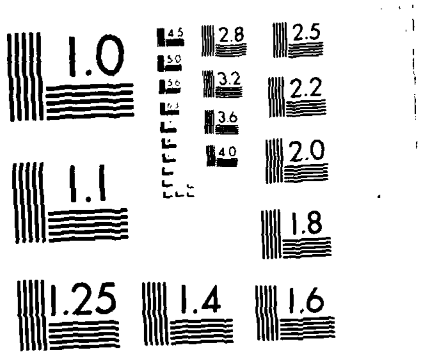
AD-A153 677 A TASK BATTERY FOR APPLIED HUMAN PERFORMANCE ASSESSMENT RESEARCH(U) AIR FORCE AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB OH C A SHINGLEDECKER NOV 84 1/1

UNCLASSIFIED AFAMRL-TR-84-071 F/G 5/8 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A153 677

(2)

AFAMRL-TR-84-071

A TASK BATTERY FOR APPLIED HUMAN PERFORMANCE ASSESSMENT RESEARCH

Clark A. Shingledecker
AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY

November 1984

Approved for public release; distribution unlimited.

DTIC FILE COPY

AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

85 04 16 125

NOTICES

TECHNICAL REVIEW AND APPROVAL

AFAMRL-TR-84-071

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Regulation 169-3.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Air Force Aerospace Medical Research Laboratory

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFAMRL-TR-84-071
6a. NAME OF PERFORMING ORGANIZATION Air Force Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable) HEG	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) Wright-Patterson AFB Ohio 45433-6573	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-82-C-0511
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
62202F	7184	14	07

11. TITLE (Include Security Classification)
See reverse side.

12. PERSONAL AUTHOR(S)
Clark A. Shingledecker

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ______ TO ______	November 1984	52

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Human Performance Task Battery
Human Performance Tests

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The Criterion Task Set (CTS) V1.0 is a battery of tests designed to place selective demands on the elementary mental resources and information processing functions of the human operator. Currently under development as a general purpose research tool for applied investigations of human performance capabilities, the CTS V1.0 includes nine standardized tasks which are implemented in user-friendly software on an inexpensive microcomputer system.

This interim report describes the theoretical basis of the CTS, the rationale used in its construction, and the parametric research conducted to establish standard loading levels, pacing rates, and training requirements for each of the tasks. Potential research applications of the CTS are discussed as well as planned developmental efforts intended to enhance its utility. Documentation for the CTS V1.0 is provided in the form of detailed individual task descriptions and brief summaries of hardware requirements and software characteristics.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Clark A. Shingledecker	22b. TELEPHONE NUMBER (Include Area Code) (513) 255-2558 22c. OFFICE SYMBOL AFAMRL/HEG

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

(11. continued)

A TASK BATTERY FOR APPLIED HUMAN
PERFORMANCE ASSESSMENT RESEARCH

PREFACE

This report documents the initial development of a human performance test battery known as the Criterion Task Set (CTS). The research effort which led to the construction of the CTS was performed in support of AFSC Project 7184, Man-Machine Integration Technology, by the Air Force Aerospace Medical Research Laboratory (AFAMRL), Human Engineering Division, Wright-Patterson Air Force Base, Ohio 45433.

The author gratefully acknowledges the technical contributions under Contract F33615-82-C-0511 of personnel from Systems Research Laboratories, Inc. (SRL), 2800 Indian Ripple Road, Dayton, Ohio 45440. Special thanks are due to Mr. Mark S. Crabtree of SRL who made the development of the CTS possible through his diligent efforts in the design of software and hardware and in the support of key experiments.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

Section 1
AN OVERVIEW OF THE CRITERION TASK SET

HUMAN PERFORMANCE ASSESSMENT

In formulating recommendations for the design and operation of systems, human factors and human performance specialists rely on several approaches to gathering information about the capacities and limitations of the human operator. Relevant knowledge often can be acquired from handbooks, data bases, and the open scientific literature. However, because the biological and cognitive sciences are not fully mature, a majority of questions about the human operator must be addressed with analytical and experimental methods which can be used to generate the required data.

Despite the recognized importance of these applied research methods, a recent report of the National Research Council Committee on Human Factors revealed that little apparent effort has been invested in their documentation or in the development of new practical techniques to meet emerging human factors needs (National Research Council, 1983). This failure to design and disseminate methods of analysis and investigation has impaired the efficiency with which human factors data are produced and subsequently utilized in system development. Lacking a variety of standardized data collection tools, human factors organizations are often forced to invest effort and financial resources in the development and validation of tailored methods for each operational problem with which they are confronted. In many cases, such additional requirements are counterproductive and may be intolerable in an applied science environment where critical system decisions must be made in a timely and cost effective fashion.

Although the National Research Council report focused on techniques used by field practitioners in specific system design activities (e.g., task analysis and checklists), similar observations may be made regarding data collection methods used in applied human performance research conducted for more general purposes. A variety of enduring research problems in human factors involve an assessment of the performance capabilities of the human operator. Specific applications of such measurements range from investigations of the effects of stressors and work environments on performance to validation of informal field methods of assessing operator capacities. Typically, these research problems require quantitative measurement of task performance under relatively well controlled testing conditions. Furthermore, a common purpose of these research efforts is to provide generalizable results by measuring performance on elementary tasks which contain the basic components of more complex operational behaviors.

The traditional method of addressing this type of human performance question is time consuming and inefficient. Commonly, several different experimental plans are pursued which reflect the diverse views and theoretical positions of individual investigators. The implicit goal of this research strategy is to provide the scientific community with ample opportunity for the development of creative approaches to the problem in the hope that a final synthesis of data occurring at some undetermined future date will offer solutions to applied problems. Regretably, such a basic science research model is of questionable value when hard decisions are required by the people responsible for deploying and operating systems.

While traditional research approaches can, and should, be defended by appealing to the fact that scientific understanding of human cognitive processes is incomplete, alternative methods are needed which provide a rapid response based on state-of-the-art knowledge. Specifically, standard techniques for assessing human performance capabilities should be developed and employed in a concerted and economical manner in order to produce a data base for answering current operational questions. The purpose of this report is to describe and document an effort to develop a standardized performance assessment methodology which is aimed at meeting the requirements outlined above.

DEVELOPMENT OF THE CRITERION TASK SET

The Criterion Task Set (CTS) is a battery of tasks which has been developed to provide an instrument for human performance assessment that is both practical and firmly based in current theoretical models of perceptual-motor and cognitive behavior. The component tasks of the CTS were designed to place selective demands on the functional information processing resources of the human operator. These elementary resources are hypothesized to be major determinants of a variety of complex task behaviors which occur in military and civilian work environments. In order to make the CTS a highly usable and valid applied research tool, investigations have been conducted at the U.S. Air Force Aerospace Medical Research Laboratory to standardize training requirements, task parameters, and loading levels. The practical utility of the CTS is further enhanced by its implementation in user-friendly, menu-driven software on an inexpensive and reliable microcomputer system.

Both analytical and empirical methods were employed in the selection and development of the CTS tasks. Initially, a model reflecting current knowledge of the human information processing system was synthesized from an extensive review of the literature on cognitive function. While a number of theoretical positions are represented in the model, primary components were derived from multiple resource theories (e.g., Wickens, 1981) and processing stage theories (e.g., Sternberg, 1969). The model was constructed to outline a minimal set of potential processing resources or capacities, and to identify a group of basic information processing activities that are performed by these structural and energetic elements. A graphic representation of the model is shown in Figure 1.

Briefly, the model hypothesizes three primary stages of processing and associated resources dedicated to perceptual input, central processing, and motor output or response activities. Within each of these stages separate resources are associated with the mode of input (visual/auditory), the code in which central processing is performed (spatial imaginal/abstract symbolic), and the mode of response output (manual/vocal). Finally, the central processing stage is divided to differentiate between working memory as a locus of central processing, and the specific types of processing or decision functions that occur at this stage. Thus, separate elements are defined to specify the basic limits of short-term storage and recall, and to identify three qualitatively different levels of central decision activity. The most elementary of these activities is information manipulation or transformation based on implicit or memorized rules. Such activity is

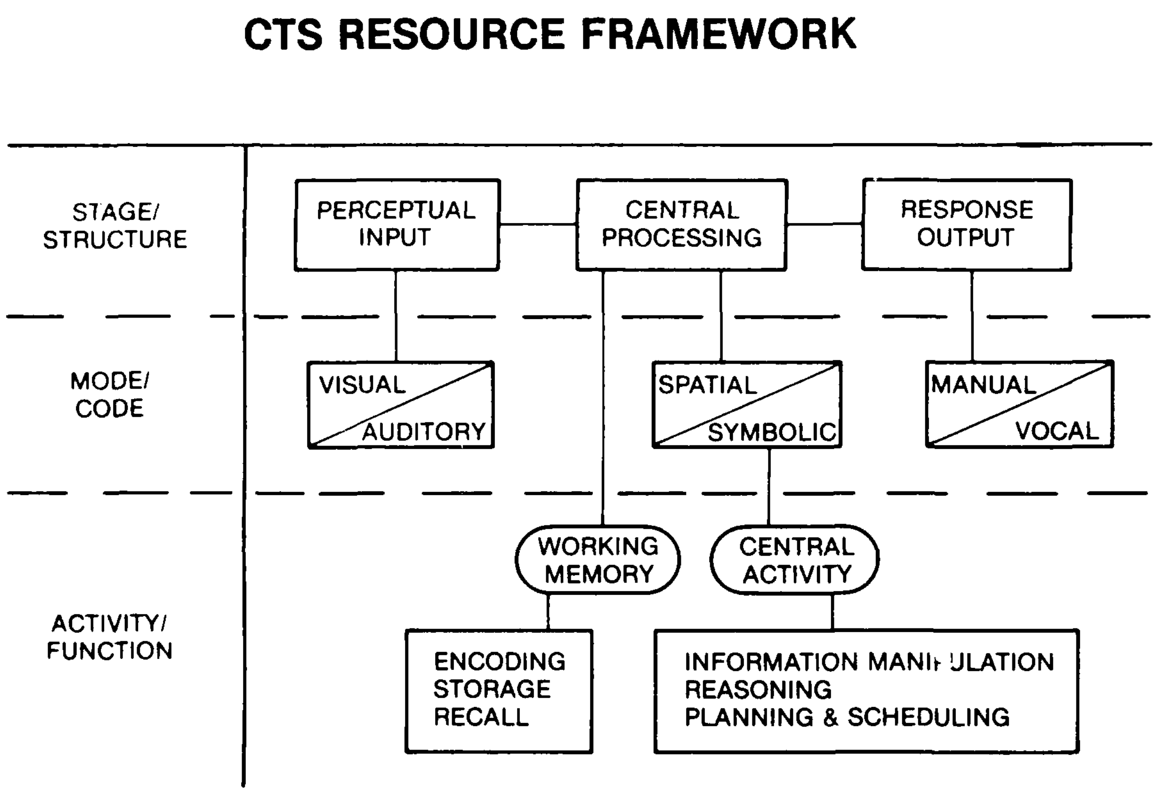


Figure 1. CTS Resource Framework

CTS CONTINUOUS RECALL DATA

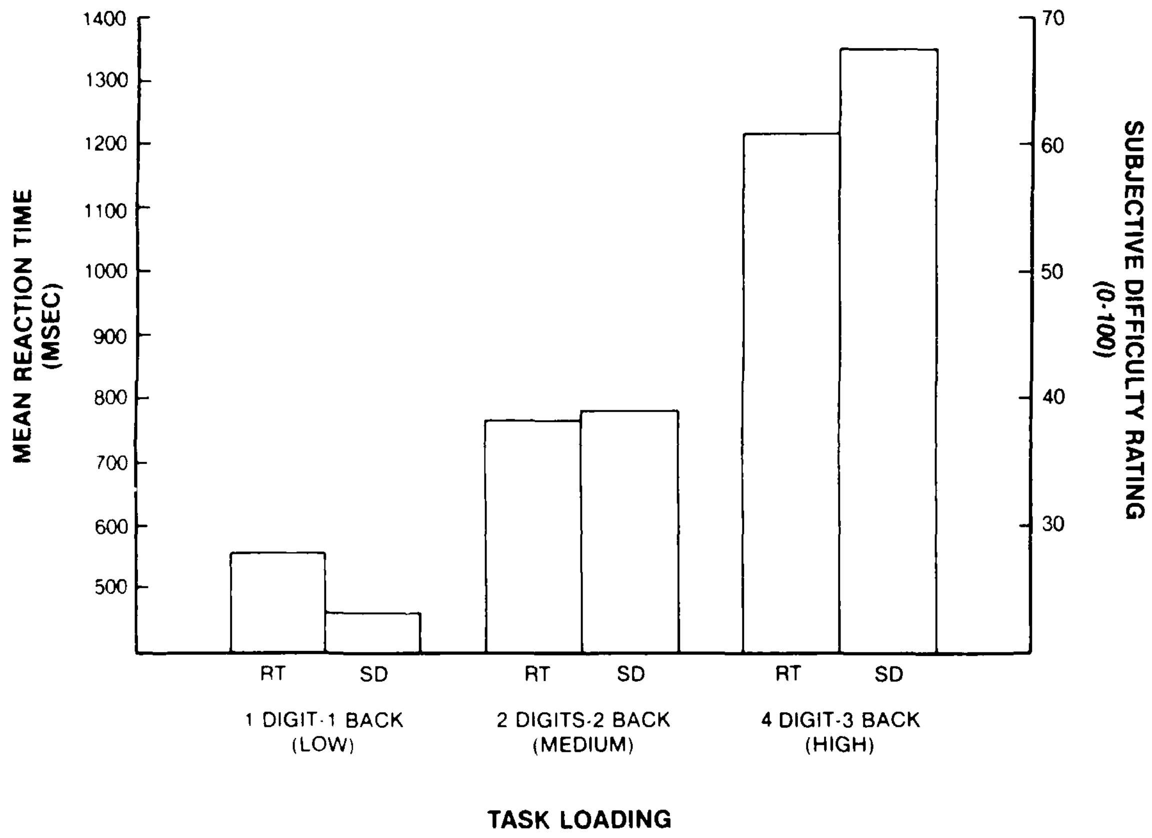


Figure 4. CTS Continuous Recall Data

information on a continuous basis. Task difficulty is manipulated by varying the number of digits which comprise each item, and the length of the series which must be maintained in memory in order to respond to recall probes.

Loading Conditions

Research conducted at the AFAMRL workload laboratory has shown that three significantly different task demand levels are produced by the following conditions: low demand--one digit per item, recalling one position back; medium demand--two digits per item, recalling two positions back; high demand--four digits per item, recalling three positions back. Mean reaction times and subjective task difficulty ratings for these conditions are shown in Figure 4.

Stimuli

Computer generated one, two, and four digit numbers are displayed serially on a CRT screen with the following restrictions: (1) test numbers must be randomly generated; (2) only the numerals 1-9 are used; (3) roughly half of the probe numbers must result in a recall comparison of "same." Test numbers and probe numbers are simultaneously presented, as well as terminated. The test numbers always appear below a line centered on the CRT while the probe numbers appear directly above the line (Figure 5).

Testing Procedure

Major practice effects are eliminated with five 3-minute trials at each loading level. However, extension of training to seven trials produces more stable performance. Subjects are encouraged to respond as rapidly and accurately as possible.

In all conditions, the task is subject paced within the limits of selected deadline reaction times. Maximum acceptable reaction time in the training mode is 15 seconds for all conditions. If the subject does not respond

Testing Procedure

Although extensive practice is not required, subjects should be familiarized with the task by passively observing signals of each probability bias pointed out by the experimenter. It is necessary for subjects to see each probability bias level before testing so that the difference in appearance of signals of each probability can be appreciated. Instructions specify that the subject not respond until he/she is certain that a signal is present. In other words, the strategy of responding more frequently than necessary to avoid missing signals is undesirable. Also, the subject is told that either two or three signals will appear in each 3-minute test period, and periods are equally likely to contain two or three signals. Regardless of the number of signals, a minimum of 25 seconds separates the offset of a signal and the onset of the next. Responses to signals are made on keys numbered and spatially arranged to correspond to the dials. Performance measures include reaction time to correctly identified signals, number of false positives (responses when no signal is present), and misses (overlooked signals).

CTS CONTINUOUS RECALL TASK

Description

The CTS Continuous Recall task is a standardized loading task designed to place variable demands upon processing resources associated with encoding and storage in working memory. The task requires an operator to utilize both immediate and short-term memory of numbers under continuously changing storage states (Hunter, 1975). The memory test consists of a random series of visual presentations of numbers which the operator must encode in a sequential fashion. As each number in the series is presented for encoding, a probe number is presented simultaneously. The operator must compare this probe number to a previously presented item at a prespecified number of positions back in the series. Once the operator has made the appropriate recall, he must decide if that item is the same as/or different from the probe number. Thus, the task exercises working memory functions by requiring operators to accurately maintain, update, and access a store of

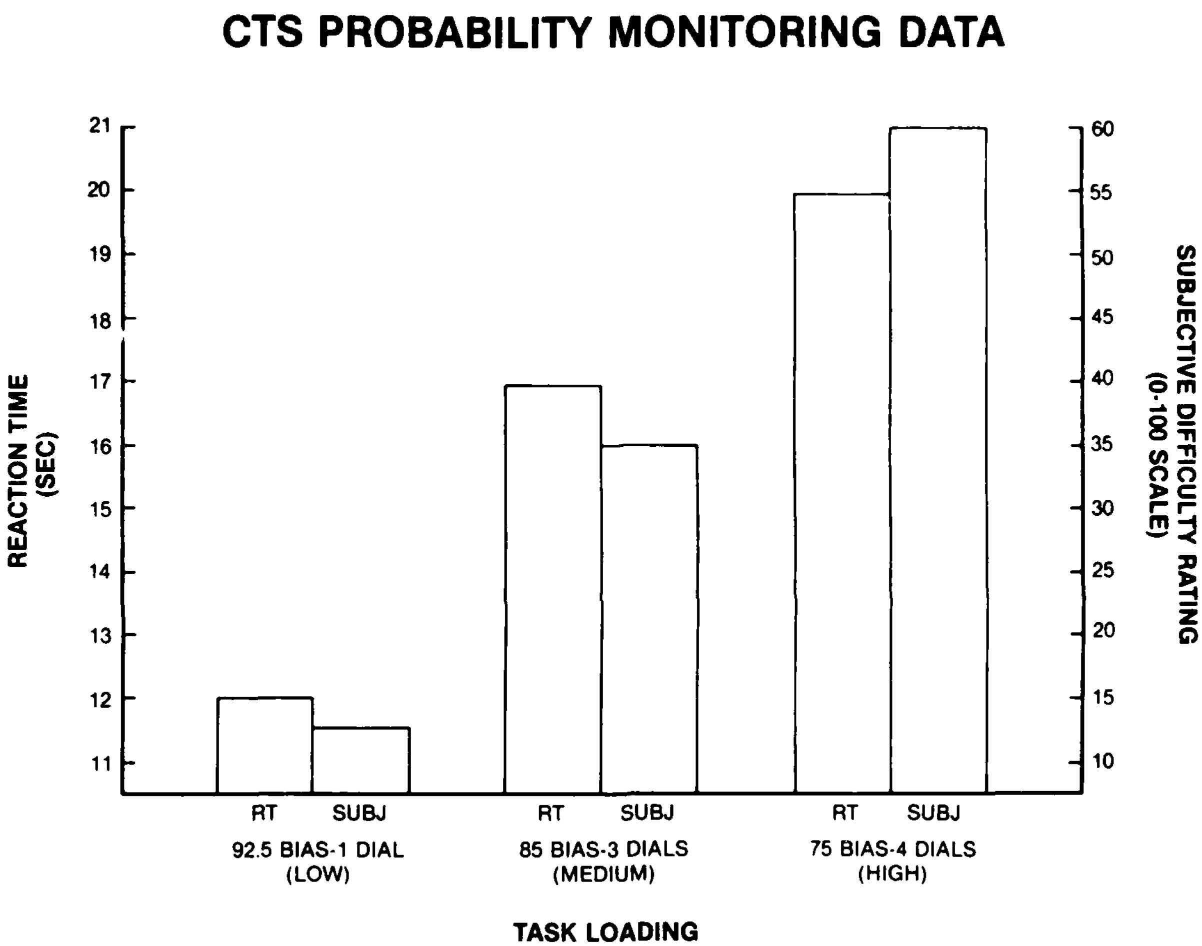


Figure 3. CTS Probability Monitoring Data

pressing an appropriate response key, biased dials can be corrected to the nonsignal (random pointer movement) state.

Nonsignal Condition

When no signal is present, the pointer moves to each position with equal probability (1/6). When more than one dial is to be monitored, the pointer movement on each dial is independent of the others. Pointer position is updated at the rate of 2 moves/sec. Dials always appear in the same screen location (i.e., dial #1 is always located at the upper-center of the screen, dial #2 at the middle-left, etc.). In the single dial condition, dial #1 is displayed; in the three dial condition, dials 1, 2, and 3 are shown; and in the four dial condition, all four dials are displayed.

Signal Condition

If undetected, a signal lasts 30 seconds and occurs over 60 pointer moves. When a signal occurs in the high discriminability condition, 57 of the 60 pointer moves appear on one side of the dial (95/5 percent probability bias); in the moderate discriminability condition, 51 of the 60 moves occur on the favored half (85/15 percent probability bias); and in the low discriminability condition, 45 of the 60 moves occur in the bias direction (75/25 percent probability bias). Within these constraints, however, pointer movement is randomly determined. Biases are equally likely to appear on either half of the displays and on any given display.

Loading Conditions

Research conducted at the AFAMRL workload laboratory has shown that three significantly different task demand levels are produced by the following task conditions: low demand--one dial at the 95/5 percent bias level; medium demand--three dials at the 85/15 percent bias; high demand--four dials at the 75/25 percent bias level. Mean reaction times and subjective task difficulty ratings for these conditions are depicted in Figure 3.

CTS PROBABILITY MONITORING TASK DISPLAY

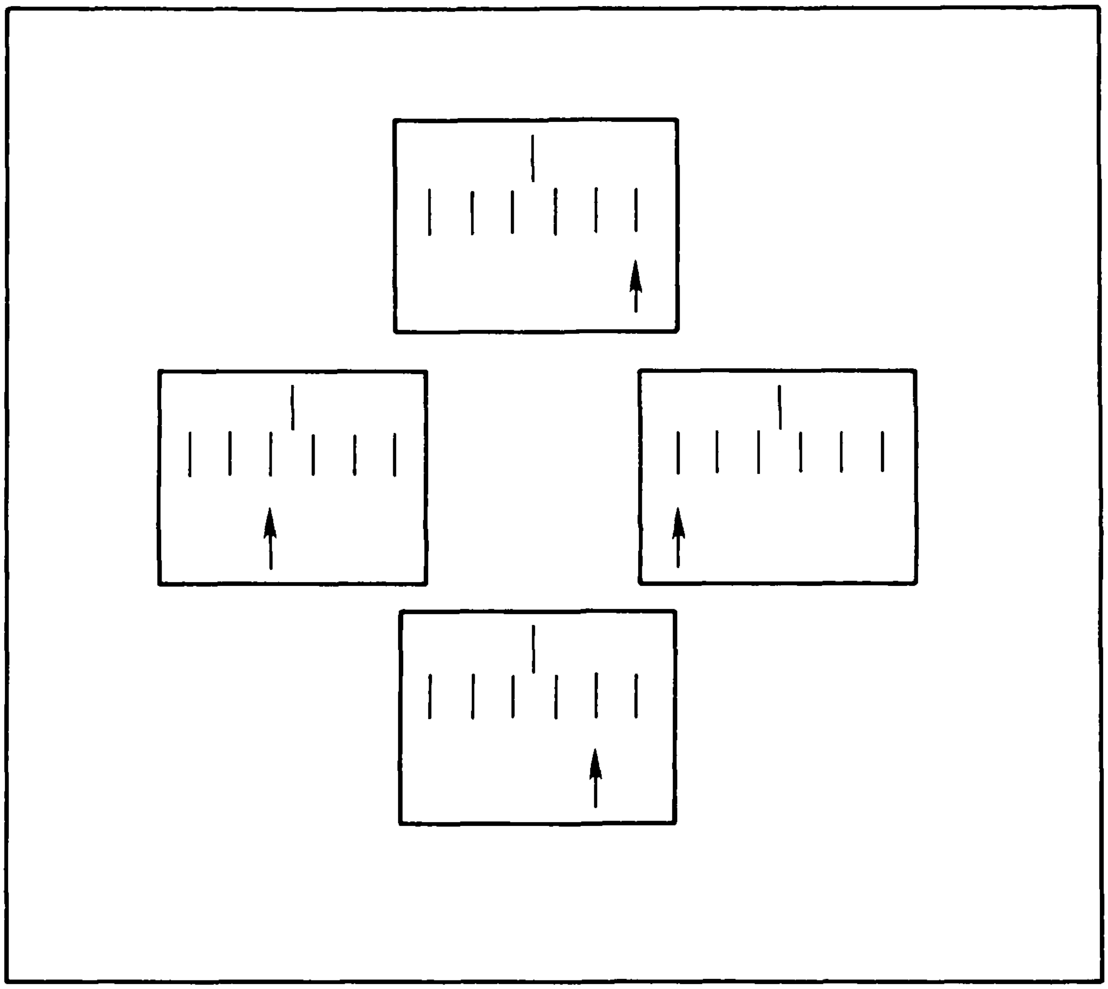


Figure 2. CTS Probability Monitoring Task Display

each loading level. If a subject responds to a stimulus prior to the deadline, a new stimulus appears immediately. However, if the subject fails to respond prior to the deadline, a new stimulus is automatically presented and the items is scored as a missed response. The deadline conditions for each loading level within each task were established by calculating mean reaction times for trained subjects and adding three standard deviations of the mean to that value.

In all CTS tasks, a single test trial at any level of loading has a standardized duration of three minutes. Software for the CTS is written in BASIC and is implemented on an inexpensive microcomputer system which is described in Appendix A of this report.

CTS PROBABILITY MONITORING

Description

The CTS Probability Monitoring task, based on a paradigm developed by Chiles, Alluisi and Adams (1968), is a standardized loading task designed to place variable demands on the visual perceptual information processing resources of the human operator. The task includes three fixed loading levels produced by variations in the number of signal sources (dials) and in the discriminability of signals. In the task, subjects are required to monitor one, three, or four computer generated displays having the appearance of electro-mechanical dials. Each display consists of a row of six vertical hashmarks with a seventh mark offset above the others to indicate the center of the dial. A number appears to the left of each dial to identify it and each dial is circumscribed by a rectangular "bezel." A representation of the display, drawn roughly to scale, is shown in Figure 2. Under normal (nonsignal) conditions a pointer located below the hashmarks moves from one position to another in random fashion to simulate the pointer fluctuations on an actual dial. At unpredictable intervals, the pointer on a display begins to move nonrandomly, staying predominantly to the left or right half of the dial. These biases in pointer movement are the targets or "signals" to which subjects are instructed to respond. By each loading level. If a subject responds to a stimulus prior to the deadline, a new stimulus appears immediately. However, if the subject fails to respond prior to the deadline, a new stimulus is automatically presented and the item is scored as a missed response. The deadline conditions for each loading level within each task were established by calculating mean reaction times for trained subjects and adding three standard deviations of the mean to that value.

In all CTS tasks, a single test trial at any level of loading has a standardized duration of three minutes. Software for the CTS is written in BASIC and is implemented on an inexpensive microcomputer system which is described in Appendix A of this report.

CTS PROBABILITY MONITORING

Description

The CTS Probability Monitoring task, based on a paradigm developed by Chiles, Alluisi and Adams (1968), is a standardized loading task designed to place variable demands on the visual perceptual information processing resources of the human operator. The task includes three fixed loading levels produced by variations in the number of signal sources (dials) and in the discriminability of signals. In the task, subjects are required to monitor one, three, or four computer generated displays having the appearance of electro-mechanical dials. Each display consists of a row of six vertical hashmarks with a seventh mark offset above the others to indicate the center of the dial. A number appears to the left of each dial to identify it and each dial is circumscribed by a rectangular "bezel." A representation of the display, drawn roughly to scale, is shown in Figure 2. Under normal (nonsignal) conditions a pointer located below the hashmarks moves from one position to another in random fashion to simulate the pointer fluctuations on an actual dial. At unpredictable intervals, the pointer on a display begins to move nonrandomly, staying predominantly to the left or right half of the dial. These biases in pointer movement are the targets or "signals" to which subjects are instructed to respond. By

Section 2
CTS TASK DESCRIPTIONS

GENERAL INFORMATION

The remainder of this report comprises a set of descriptive summaries for the nine tasks included in the CTS V1.0. Each description briefly defines the function or resource tested by the task as well as loading and pacing parameters. Training requirements are also provided along with any special procedural or instructional information needed to conduct testing sessions. Finally, where appropriate, representative data from trained subjects are summarized.

In general, each CTS task includes three conditions which can be selected by the experimenter to produce low, moderate, and high levels of task demand. The single exception is the CTS Interval Production Task which assesses response timing capability and provides a single testing condition. As noted previously in this report, the selection of standard loading levels for each task was based on data acquired from developmental studies in which predicted loading variables were parametrically manipulated. As a result, it should be noted that the standard loading levels for each task must be interpreted as valid differences at an ordinal level of measurement. That is, no rigorous inferences can be made concerning the absolute magnitude of the differences in task demand between low and moderate loads or moderate and high loads. In addition, for both theoretical and statistical reasons, common loading levels on different tasks should not be interpreted as being equated on any scale of measurement.

The CTS was designed to place highly selective demands on individual mental functions. Since time pressure is a generalized loading factor which would affect workload in any of these functions, task pacing was not used to produce explicit variations in the demand of the CTS tasks. Thus, training on all tasks is conducted under essentially subject-paced conditions. Test trials are also subject paced but impose mildly restrictive time limits on the subjects' response in order to maintain trained performance levels. In all discrete stimulus tasks of the CTS, a response deadline is defined for level of demand. Furthermore, the availability of ordered levels of task difficulty enables scaling of the magnitude of stress effects without resorting to strong assumptions regarding the nature or generalizability of each measure of performance contained in the battery. For example, it is possible to express the effects of a stressor on a specific information processing resource or function in terms of the number of statistically significant directional changes in performance observed in its corresponding CTS task when tested across all three loading levels under baseline and treatment conditions.

In its present form, the CTS is expected to serve as a useful applied research tool for the topic areas described above. However, continued developmental research is necessary to expand the theoretical foundation and practical utility of the CTS. Basic validation studies will be performed with the CTS using both mutual task interference and multivariate, factor analytical methodologies. These diverse approaches to investigating underlying components of human performance will be employed to explore the structural independence of the CTS tasks and may result in future modifications of the battery. A second research effort will attempt to devise a modeling and task analytical system which would permit decomposition of operational tasks to the human performance model embodied in the CTS. Since the models used by task analysts and human performance researchers rarely coincide, generalizations made about real-world operational performance effects based on those obtained from elementary human performance tests often have doubtful validity. The purpose of this research will be to construct an explicit methodology linking the CTS tasks to corresponding operational behaviors. Successful completion of the effort will enhance the utility of the CTS by making it possible to formulate specific quantitative predictions about complex task performance effects from those observed in the CTS under treatment conditions.

physiological indices of workload. Preliminary research in which a subset of the CTS tasks were used to evaluate a secondary task index of workload showed that the CTS can be employed to diagnose the specific types of information processing demands to which a particular metric is optimally sensitive (Shingledecker, Acton, and Crabtree, 1983). The goal of this continuing effort is to provide human factors specialists and designers with prescriptive guidelines for the use of workload assessment techniques.

A second broad area of investigation to which the CTS can be applied as a standardized test instrument is the assessment of human performance capabilities. When used for this purpose, the tasks comprising the CTS may be employed in a diagnostic fashion to measure and predict the effects of extreme environments and biochemically active agents on human performance. Although numerous single tasks and task batteries have been developed for such stress research in the past, the design features of the CTS make it particularly amenable to this type of application.

Unlike many other performance tests, the CTS tasks are linked to a specified model of human performance which identifies generic resources and mental activities responsible for a wide range of operator behaviors. This model provides an interpretive framework for the observed performance effects of a treatment condition. Thus, rather than merely demonstrating that one or more performance measures are altered following exposure to a particular stressor, the pattern of results obtained on the CTS should be a sensitive indicator of the specific type of information processing function affected, and of the general classes of tasks which would be expected to show deterioration.

The manipulable task demand characteristics of the CTS also enhance its utility for general human performance assessment. Currently, eight of the nine tasks in the CTS offer the capability to test corresponding perceptual, central, and response resources at empirically selected low, moderate, and high levels of loading. This feature permits the analysis of interactions between stress variables and task difficulty. Consequently, it is likely that the CTS will be more sensitive to the effects of a stressor with unknown strength than other tasks which test performance at only a single

TABLE 1. CTS V1.0 TASKS

CTS Task	Tested Resource/Function	Typical Resource Based Behaviors
Probability Monitoring	Visual Perceptual Input	Scanning Detection Monitoring
Continuous Recall	Working Memory Encoding	Memorizing Keeping Track of Events
Memory Search	Working Memory Retrieval	Recalling Recent Events
Linguistic Processing	Symbolic Information Manipulation	Analysis of Meaning Language Comprehension Classification of Events
Mathematical Processing	Symbolic Information Manipulation	Computing Calculating Comparison of Values
Spatial Processing	Spatial Information Manipulation	Maintaining Orientation Identifying Patterns Analyzing Position
Grammatical Reasoning	Reasoning	Problem Solving Analyzing Relationships Logical Thinking
Unstable Tracking	Manual Response Speed and Accuracy	Continuous Control Error Correction Control Actuation
Interval Production	Manual Response Timing	Scheduling Movements Coordinating Sequential Responses

CTS APPLICATIONS AND RESEARCH

The theoretical basis and standardized features of the CTS make it potentially applicable to a number of research problems in the areas of human performance assessment and human factors. One of these problems for which the CTS was originally designed is the comparative evaluation of measures of mental workload. In this application, the individual components of the CTS are being used as primary loading tasks to assess the reliability, sensitivity and intrusiveness of a number of proposed behavioral, subjective, and for the CTS. The tasks were then subjected to extensive empirical evaluation in order to determine training times needed to attain stable performance levels, to fix task pacing rates, and to establish standard task loading levels. In each of these studies predicted task loading parameters were manipulated factorially using a within-subjects repeated measures experimental design ($N > 6$). Training in all conditions extended over ten 1-hour sessions conducted on successive working days. Performance was scored on appropriate speed and accuracy measures for each task, and training asymptotes were identified using successive analysis of variance methods. Multiple comparisons performed on post-asymptotic performance indices were used to select three loading levels for each of the individual tasks. These levels were corroborated by similar analyses of subject ratings of task difficulty and complexity.

In order to assess the reliability of the findings obtained in the individual task studies described above, a combined replication study was also performed. Twenty AFROTC cadets were trained on all tasks of the CTS. The results confirmed that the originally selected task parameters produced reliable loading effects after prescribed training had been completed. Detailed descriptions of this replication study and of the original parametric task evaluations will appear in future reports. Because further developmental research is in progress, the current battery represents a first version of the CTS. The nine tasks included in the CTS V1.0 are listed in Table 1. Additional tasks which will be retained in the CTS and will appear in a second version are the CTS Auditory Monitoring Task and the CTS Supervisory Control Task. These tasks are intended to assess auditory perceptual processes and planning and scheduling decision activity, respectively.

common in language comprehension, mathematical computation, and pattern analysis. Reasoning functions represent a higher level of central activity which involve the extraction of relational rules from the information presented. This activity occurs during problem solving and in tasks requiring logical analysis. Finally, planning and scheduling central processing activities are characterized by multiattribute decision requirements where several factors must be considered in selecting an optimal course of action. These activities are typical in complex system supervision and planning tasks.

It should be noted that no universal concensus presently exists regarding an appropriate theory of human performance, and that the model outlined above was intended to act only as a descriptive summary of state-of-the-art research findings and conceptual approaches. The primary purpose for constructing the model was to guide the development of a set of representative tasks for the CTS. This was accomplished by operationally defining the model in terms of the characteristics of tasks which would place predominant demands on each of its represented elements. For example, visual perceptual resources were defined in terms of a task which would require detection of simple signals embedded in "noise" as well as monitoring and scanning behaviors. Thus, loading factors such as stimulus discriminability and numerosity of display sources were used as criteria for task selection. At the same time, minimal central processing and response demands for the task were considered essential so that loading would be confined primarily to the input stage of processing.

General task descriptions similar to that summarized above were generated for the remaining resources and functions included in the model. In each case the ability to manipulate task demand and to minimize loading on resources not tested by the task were used as major selection criteria. Additionally, where possible, an attempt was made to achieve face validity in the CTS tasks in order to enhance subject acceptance and the ease with which test results could be generalized to actual operational tasks.

Iterative screening of a wide range of tasks from the literature on cognitive and psychomotor performance resulted in the selection of 11 tasks

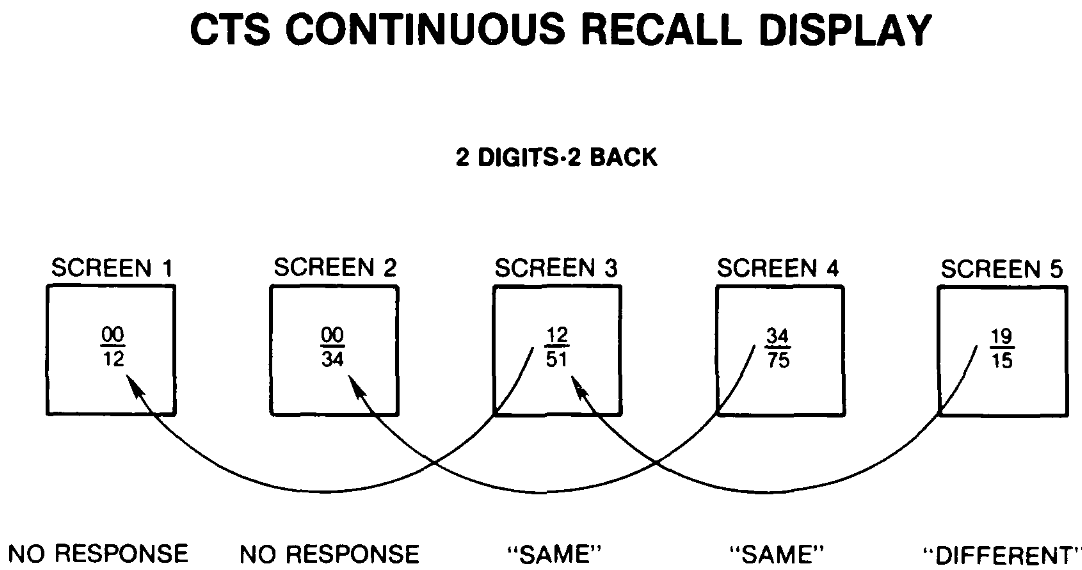


Figure 5. CTS Continuous Recall Display

within 15 seconds after the onset of a test item, the next item is automatically presented.

In the testing mode the reaction time deadlines are reduced: 1.1 seconds for the one digit, 1 back condition; 1.7 seconds for two digits, 2 back; and 2.3 seconds for four digits, 3 back. The numbers display is approximately 1.25 inches high, each number is approximately .25 inch x .13 inch, and should be viewed from a distance of roughly 60 cm. Responses are entered on a two button keypad. Subjects are given feedback concerning the accuracy of their performance after each 3-minute test period to ensure that an acceptable speed-accuracy trade-off is maintained.

CTS MEMORY SEARCH TASK

Description

The CTS Memory Search task, based on Sternberg's (1969) memory search paradigm, is a standardized task designed to place variable demands on human information processing resources dedicated to short-term memory retrieval functions. In the memory search task, a small set of items (the "memory set") is first presented to the subject for memorization. A series of test items are then presented to the subject one at a time, and the subject must respond positively if the test item was contained in the memory set, or negatively if not. Reaction time is measured from the onset of the test item to the response. The CTS version of the task is composed of three fixed demand levels produced by variations in the number of items to be memorized. Research conducted at the AFAMRL workload laboratory has demonstrated that memory set sizes of one, four, and six items produce reliably different levels of performance and subjective workload. Mean reaction times and subjective workload ratings for these conditions are shown in Figure 6.

Stimuli

Stimulus items in the CTS memory search task are visually presented alphabetic characters. Due to the acoustic confusability of certain letters

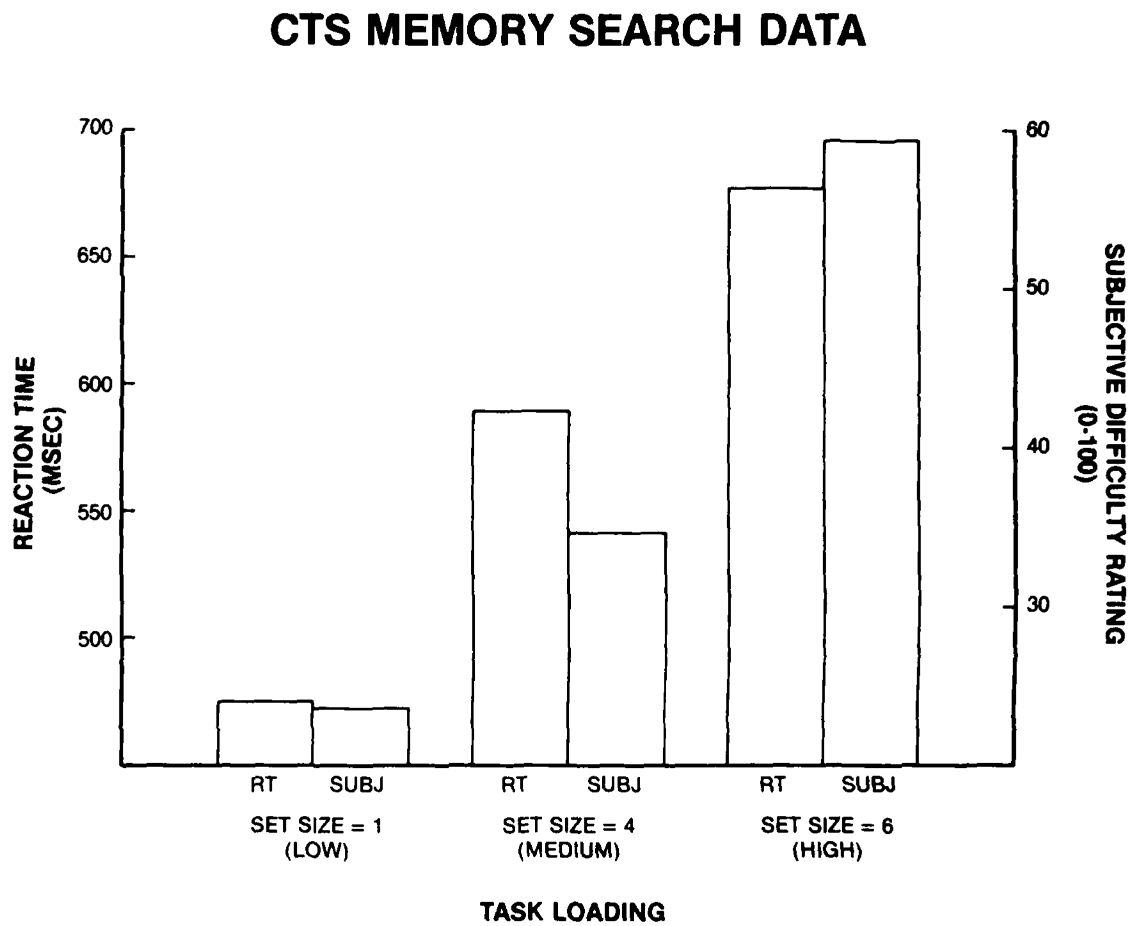


Figure 6. CTS Memory Search Data

only 17 of the 26 letters of the alphabet are used in the task (ABCEFGHIJLOQRSXYZ). Memory set items are randomly selected from the letter population, and the remaining items are used in the negative set. A new memory set is selected at the beginning of each 3-minute test period. Test items are also randomly generated with the restriction that positive and negative set items are drawn with equal probability.

Testing Procedure

Major practice effects are eliminated with seven training trials at each loading level. However, extension of training to 16 trials produces more stable performance on the memory search task. Subjects are encouraged to respond as rapidly and accurately as possible. In all conditions, the task is subject paced with a deadline, allowing the subjects to pace themselves within experimenter determined time constraints. Maximum acceptable reaction times in the training mode is 15 seconds for all memory set sizes. If the subject does not respond within 15 seconds after the onset of a test item the next item is automatically presented. In the testing mode, reaction time deadlines are reduced: 1.5 seconds for memory set size one, 2.0 seconds for set size four, and 2.5 seconds for set size six. Letters are approximately .5 x .7 cm and should be viewed from a distance of roughly 60 cm. Responses are entered on appropriately labeled keys. Subjects are given feedback concerning the accuracy of their performance after each test period to ensure that an acceptable speed-accuracy trade-off is maintained (less than 5 percent error).

CTS LINGUISTIC PROCESSING TASK

Description

The CTS Linguistic Processing task is a standardized loading task that places variable demands upon mental resources associated with the manipulation and comparison of linguistic information. The task is actually a synthesis of two tasks which have appeared in the psychological literature, the Posner letter matching task (Posner, 1967), and variations on a generic "depth of processing" task (e.g., Shulman, 1974; Craik and Tulving, 1975). The CTS Linguistic Processing task requires classification of letter and word pairs as "same" or "different" on the basis of three stimulus dimensions. Task difficulty is determined by the dimension along which stimuli are compared. Letter or word pairs are presented on a CRT. Subjects respond positively if the items match on the dimension in question or negatively otherwise. Past research with this type of task has demonstrated that response latencies and the degree of incidental learning of stimuli are influenced by the type of classification rule employed.

Loading Conditions

Experiments have been conducted at the AFAMRL workload laboratory to determine significantly different demand conditions for the task. The results indicated that three distinct levels of task demand are imposed by the following classification rules: physical letter match, in which letter pairs must be physically identical to match (low demand); category match, requiring that both letters be either consonants or vowels (moderate demand); and antonym match, in which only words opposite in meaning constitute a match (high demand). Figure 7 shows mean reaction times and subjective difficulty ratings for the three loading conditions.

Testing Procedure

Major practice effects are eliminated with five trials at each loading level. However, extension to ten trials produces more stable performance. The task is performed in 3-minute trials. A maximum response time, or "deadline," is imposed in each condition. Stimuli are displayed until the subject responds or until the deadline is reached, thus allowing subjects to pace themselves within the restriction imposed by the deadline. During training, the deadline is set at 15.0 seconds for all conditions. More restrictive deadlines are used on testing trials. For the physical identity match condition, the testing deadline is 1.0 second; for the category match condition, 1.5 seconds; and for the antonym match condition, 2.5 seconds. Letters are approximately .5 x .7 cm and are viewed from a distance of roughly 62 cm. Positive and negative responses are entered on appropriately

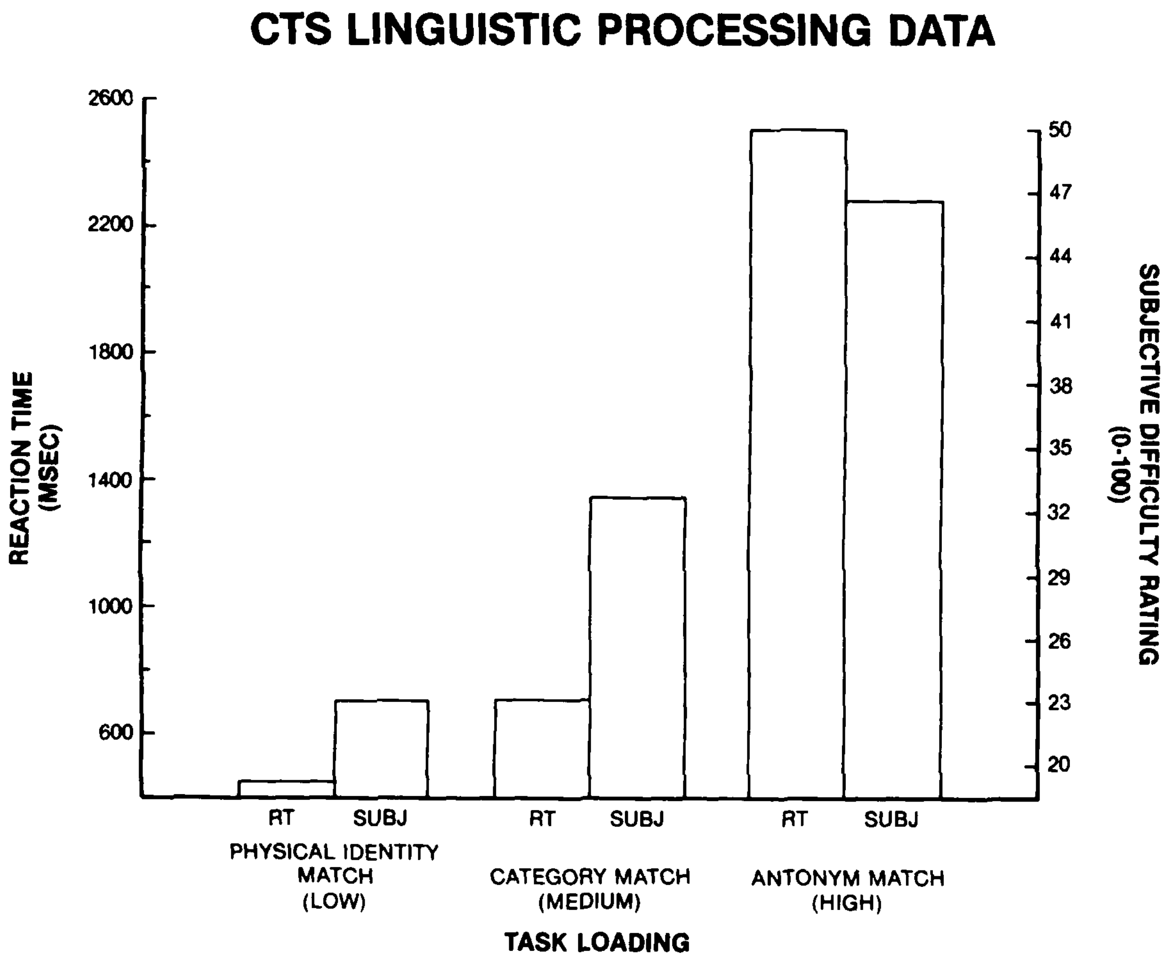


Figure 7. CTS Linguistic Processing Data

labeled keys. Subjects are instructed to respond as quickly as possible without error, and are given performance feedback after each trial to ensure that acceptable error levels are maintained. Measures of both speed (reaction time) and accuracy (percent error) are obtained.

Stimulus Generation

Letter pairs for the physical identity match rule are drawn from the population of all possible (64) combinations of both upper and lower case versions of the letters A, B, C, and E. Positive and negative letter pairs are randomly generated with equal probability. Antonyms were taken from Roget's Thesaurus (Roget, 1923). To minimize the recognition of repeated antonyms, individual words composing the antonyms are paired with both matching and nonmatching words throughout testing.

CTS MATHEMATICAL PROCESSING TASK

Description

The CTS Mathematical Processing task is a standardized loading task designed to place variable demands upon information processing resources associated with the manipulation and comparison of numeric stimuli. The task requires the subject to perform one or more simple arithmetic operations on visually presented single digit numbers to determine whether the correct answer is greater or less than a prespecified value (5). Task complexity is determined by the number and combination of operations in the problems. Research at the AFAMRL workload laboratory has shown that three significantly different task demand levels are produced by the following conditions: one-operator problems involving either addition or subtraction (low demand); two-operator problems with +-, -+, and -- operator combinations (moderate demand); and three-operator problems with ++-, +-- and -+- operator combinations (high demand). Mean reaction times and subjective task difficulty ratings for these conditions are shown in Figure 8.

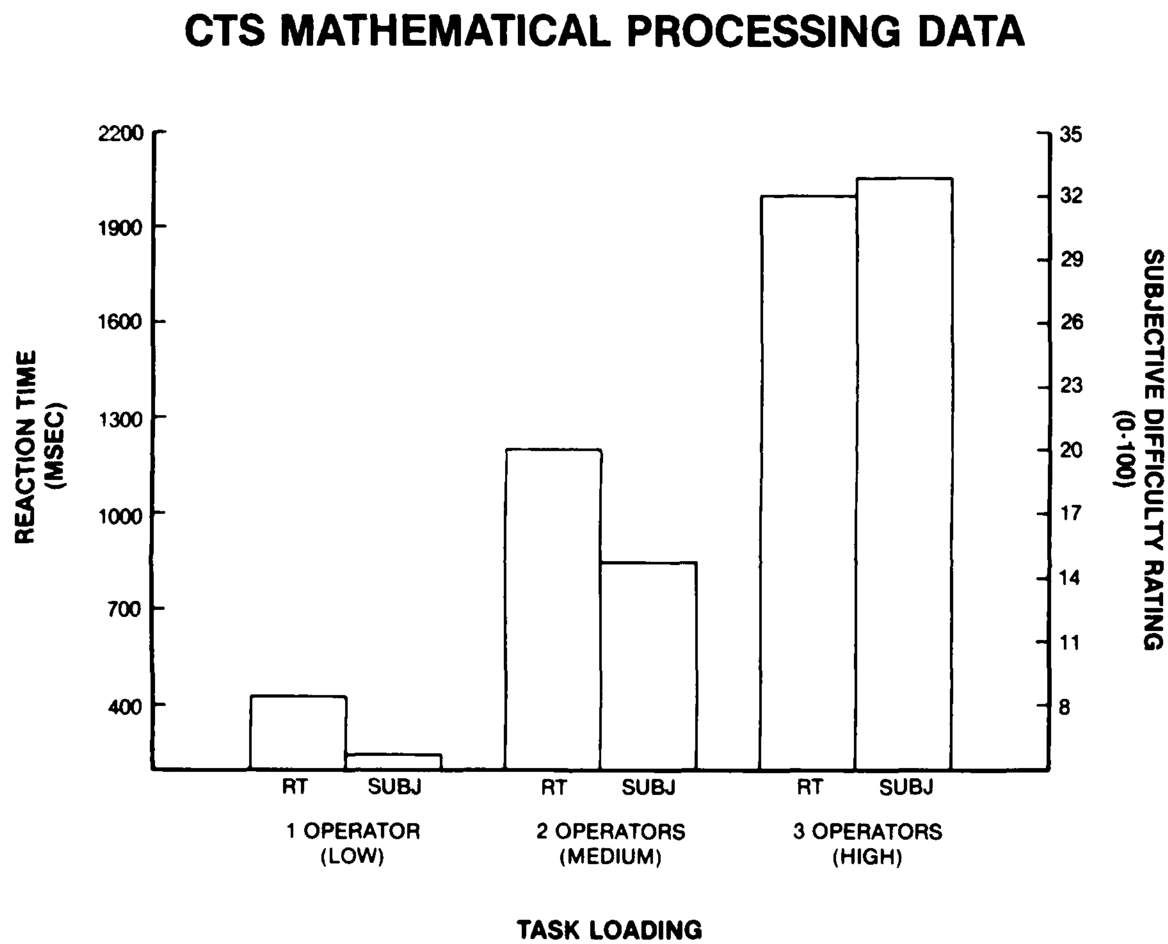


Figure 8. CTS Mathematical Processing Data

Stimuli

Math problems requiring simple addition and subtraction are randomly generated with the following restrictions: (1) only numbers 1 to 9 may be used in the problems; (2) the correct answer may be any number from 1 to 9 except 5; (3) roughly half of the problems must have an answer greater than 5; (4) when problems are solved from left to right, cumulative intermediate totals must be positive; (5) successively presented problems never have the same combination of numbers and operations in the same order and are, therefore, never identical.

Testing Procedure

The amount of practice required to reduce the effect of training to non-significant levels is dependent on the number of operators in the problem. One-operator problems require seven training trials while two- and three-operator problems require 10 training trials. Performance stability is enhanced if practice is extended to 14 and 30 trials, respectively, on the two- and three-operator conditions. Subjects should be instructed to perform the operations from left to right in order to avoid calculations with negative numbers. Problems are painted on a CRT screen from left to right at a rate that is slow enough to suggest this strategy, but not so slow as to force it (paint time = 50 msec/character). Response deadlines of varying length are imposed so that the subject can pace himself or herself within certain experimenter determined limits. If a response is not made within the deadline time, the stimulus is erased and a new one presented. On training trials, the response deadline is set at 15 seconds. During actual testing, the deadlines become more restrictive. For one-operator problems, a deadline of 1.5 seconds is imposed; for two-operator problems, the response deadline is 3.0 seconds; and for three-operator problems, the deadline is 4.0 seconds. Testing is accomplished in 3-minute trials. Subjects respond by pressing one of two appropriately labeled keys. Measures of reaction time and percent correct are taken, and subjects are given performance feedback after each 3-minute test period to ensure that an acceptable speed-accuracy trade-off is maintained.

CTS SPATIAL PROCESSING TASK

Description

The CTS Spatial Processing task is a standardized loading task designed to place variable demands upon information processing resources required for the manipulation and comparison of spatial information. Based on a task developed by Chiles, Alluisi, and Adams (1968), this CTS task requires the operator to view a series of histograms presented one at a time. The operator must determine whether the second histogram in each set of two (the "comparison" item) is identical to the first (the "target" item) and respond either positively or negatively. Target and comparison histograms are marked with the numbers 1 and 2, respectively, so that subjects can keep track. Task demands are manipulated by varying the number of bars in the histograms and the spatial orientation of the comparison histogram. Evaluation research conducted at the AFAMRL workload laboratory suggests that both the number of bars in the histograms and the orientation of the comparison histogram are effective parameters for the manipulation of task demands. Low demands are placed on the operator when two bar histograms are presented with comparison items in the 0-degree orientation. Four bar stimulus pairs with comparison items at the 90-degree and 270-degree orientations represent a moderate loading condition. Finally, six bar comparison histograms presented at the 180-degree orientation impose relatively high demands on the operator. Mean reaction times and subjective task difficulty ratings for these conditions are shown in Figure 9.

Stimuli

Computer-generated two, four, and six bar histograms are displayed one at a time on a CRT screen. Bar heights vary from 1 to 6 arbitrary units. No two bars in a histogram are the same height, and any of the six bar heights may appear regardless of the number of bars in the histogram. The first histogram in each pair is always presented in a vertical orientation, with a horizontal line under the figure and the number 1 underneath. Comparison histograms are presented at 0-degree, 90-degree, 180-degree, and 270-degree

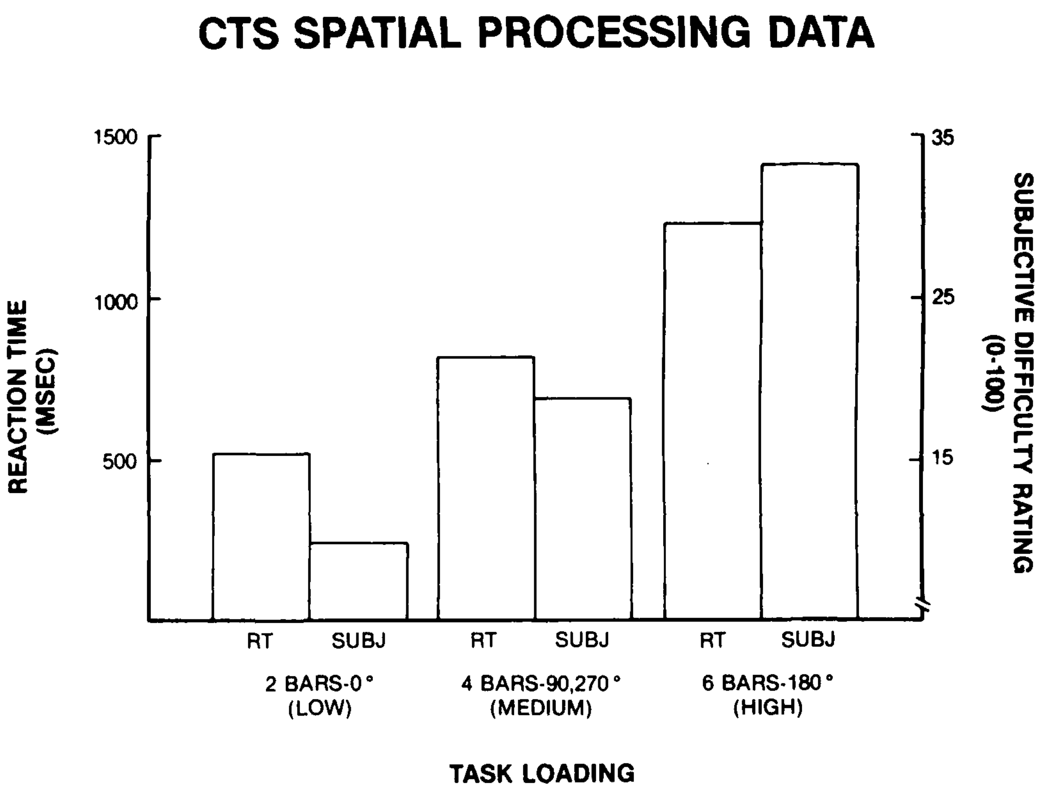


Figure 9. CTS Spatial Processing Data

orientations, and are also underlined and accompanied by a "2" to distinguish them. Target histograms are randomly selected from all possible combinations of bar heights given the number of bars in the histogram. Generation of comparison items is also random with the restriction that roughly half of the comparisons are identical to the target. When displayed on a CRT, one unit of bar height is approximately .85 cm. The tallest bar (6 units) is, therefore, about 5.1 cm high. Bars are roughly .5 cm wide and are separated by .4 cm spaces. Average viewing distance is 60 cm.

Testing Procedure

Major practice effects are eliminated with six training trials at each loading level. However, extension of training to 10 trials produces more stable performance. The task is performed in 3-minute trials. Targets are displayed for 3 seconds, followed by a short pause. Comparison histograms are displayed for a maximum of 1.5 seconds in the two bar condition, 2.5 seconds in the four bar condition, and 3.5 seconds in the six bar condition. If responses are entered before these deadlines are reached, the screen goes blank for the remainder of the deadline period. Only responses made during the interval between comparison onset and the end of the deadline period are accepted. Positive and negative responses are input on two appropriately labeled keys on a four button keypad. Measures of both speed and accuracy of responses are taken, and subjects are given feedback after each trial to ensure that acceptable accuracy levels are maintained.

CTS GRAMMATICAL REASONING TASK

Description

The CTS Grammatical Reasoning task is designed to impose variable processing demands on resources required for logical thought. The logical system contained within English grammar is used to test the ability to extract relational rules from sentence stimuli. The task derived from Baddeley's (1968) Grammatical Reasoning task. Stimulus items are sentences of varying syntactic structure accompanied by set of symbols presented simultaneously. The sentences must be analyzed to determine whether they correctly describe the ordering of the symbols in the symbol set. Task demand is influenced by the amount and complexity of grammatical analysis. Testing conducted at the AFAMRL workload laboratory has demonstrated that three different levels of grammatical demands are imposed by the following task conditions: (1) single-sentence items of variable syntactic construction describing the order of pairs of letters (all possible stimuli in the Baddeley version)--low demand; (2) items composed of two sentences worded actively and positively, and describing the positions of three symbols--moderate demand; and (3) two-sentence items worded either actively/negatively or passively/negatively and describing three symbols--high demand. Figure 10 shows mean reaction times and subjective difficulty ratings for these conditions.

Stimuli

The stimulus population for single-sentence problems is comprised of all possible combinations (32) of the following five binary conditions: (1) active versus passive wording of sentences; (2) positive versus negative wording; (3) keyword "follows" versus "precedes"; (4) order of the two symbols in the sentence; and (5) order of symbols in the symbol set. Example one-sentence items are shown in Table 2. For one-sentence (simple) items, the subject's task is to decide whether the symbol set is ordered as the sentence indicates.

TABLE 2. GRAMMATICAL REASONING, ONE-SENTENCE ITEM EXAMPLES

1.	@ precedes *	*@	(active/positive; false)
2.	@ does not follow *	@*	(active/negative; true)
3.	@ is followed by *	*@	(passive/positive; false)
4.	* is not preceded by @	*@	(passive/negative; true)

In the task conditions using two sentences (medium and high demand conditions), the object is to determine whether both sentences match in their correctness. If both sentences correctly describe the ordering of the three symbols, or if neither is correct, the subject responds positively. If one

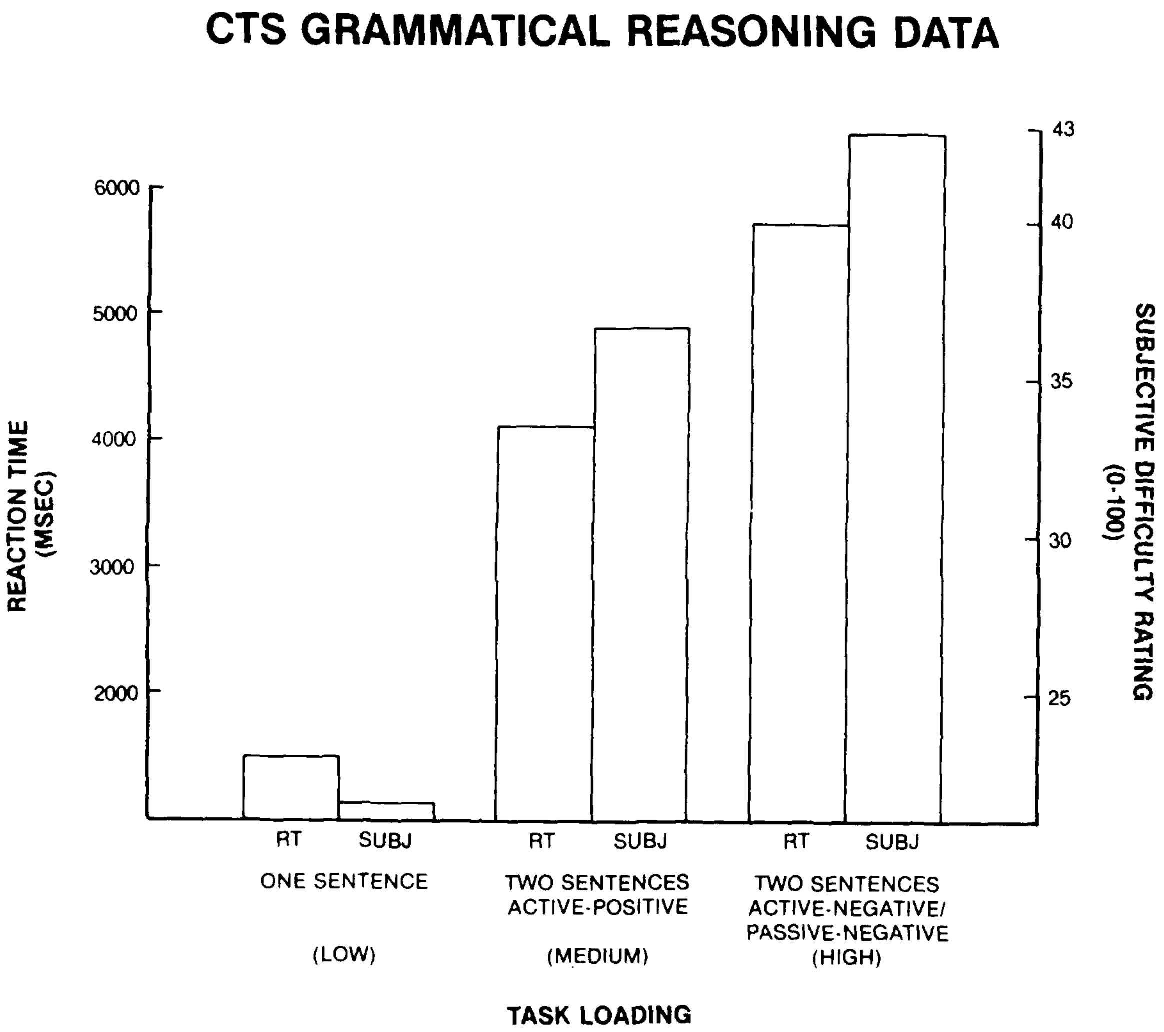


Figure 10. CTS Grammatical Reasoning Data

sentence is correct but the other not, a negative response is given. Sentences always describe adjacent symbol pairs, and are of the same grammatical form (e.g., an active/negative sentence is never paired with a passive/negative sentence). To help equate all conditions, problem sets of 32 (the total number of single sentence problems) were randomly selected for the two sentence conditions with two restrictions. First, when correctly solved, half of the two sentence problems result in a positive response. Second, combinations of sentence answers (e.g., sentence one true, sentence two true; sentence one true, sentence two false, etc.) occur equally often. Equal numbers of active/negative and passive/negative items are used in the high demand condition. Sample two-sentence items are shown in Table 3.

TABLE 3. GRAMMATICAL REASONING, TWO-SENTENCE ITEM EXAMPLES

1.	@ precedes * * follows #	@*#	(active/positive; nonmatch-false)
2.	* follows @ @ follows #	*@#	(active/positive; match-true)
3.	# does not precede * # does not follow @	*#@	(active/negative; match-true)
4.	* does not follow # # does not precede @	*#@	(active/negative; nonmatch-false)
5.	@ is not followed by # @ is not preceded by *	#@@	(passive/negative; match-true)

Testing Procedure

Major practice effects are eliminated with nine training trials at each loading level. During training, grammatical items are subject paced with a 15-second deadline for all three demand levels. If the subject does not respond within 15 seconds of the onset of the stimulus, the display is cleared and a new item is presented. During testing, deadlines vary by task condition. The deadline for the low demand condition (simple one-sentence items) is 2.5 seconds, for the moderate demand condition (two sentences, active/positive wording), 6.5 seconds; and for the high demand condition (two sentences, active/negative or passive/negative wording), 7.5 seconds. Binary responses are entered manually on appropriately labeled keys. Subjects should receive performance feedback to maintain acceptable performance levels.

CTS UNSTABLE TRACKING TASK

Description

The CTS Unstable Tracking task, similar to the critically unstable tracking task developed by Jex, McDonnell, and Phatak (1966), is a standardized loading task designed to place variable demands on human information processing resources dedicated to the execution of rapid and accurate manual responses. In the task, subjects view a video screen displaying a fixed target area centered on the screen. A cursor moves vertically from the center of the screen, and the operator attempts to keep the cursor centered over the target area by rotary movements on a control knob. The system represented by the task is an inherently unstable one. The operator's input introduces error which is magnified by the system with the result that it becomes increasingly necessary to respond to the velocity of the cursor movement as well as cursor position. Research conducted at the AFAMRL workload laboratory has demonstrated that, based on measures of tracking performance (average absolute tracking error and number of control losses) and subjective task difficulty ratings, three reliably different demand levels are produced by lambda values of 1.0 (low demand), 3.0 (moderate demand), and 5.0 (high demand). Integrated tracking error scores and subjective ratings for these task conditions are shown in Figure 11.

System Dynamics

The unstable plant dynamics of the task are a first-order divergent element of the form:

$$P_{(s)} = \frac{\lambda}{s - \lambda} e^{-ts}$$

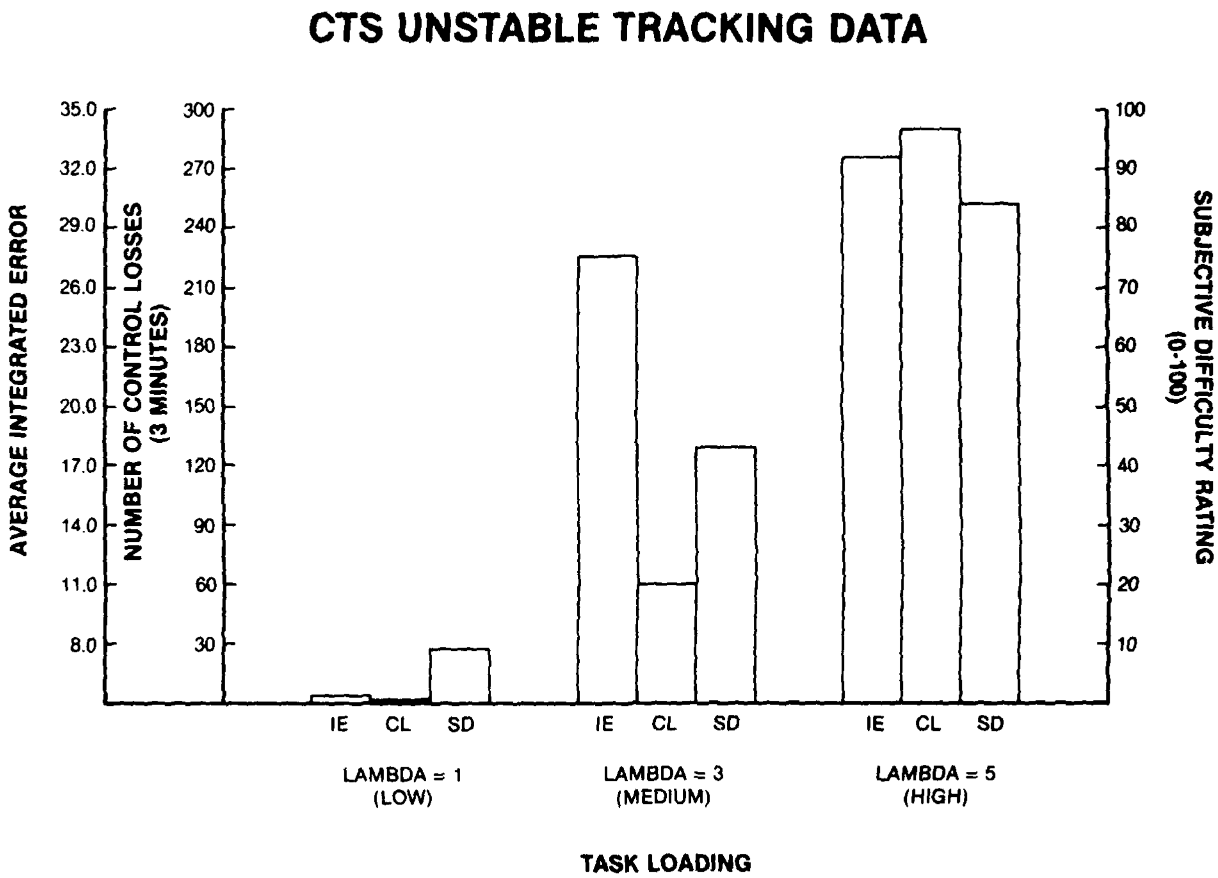


Figure 11. CTS Unstable Tracking Data

where λ (lambda) is selected by the experimenter to vary the manual control workload. The time delay term (t) in the above equation was not explicitly specified to be part of the desired dynamics, but is present in any digital implementation of a tracking loop. The magnitude of this delay was determined analytically to be no greater than 49 msec. It includes the 21 msec frame time (1000 msec/47 Hz), an 11 msec sample-and-hold (0.5 x frame time) associated with display generation, and a 17 msec sample-and-hold associated with the television frame time.

The K_1 gain term represents the hardware characteristics of the Commodore 1702 monitor [one unit of commanded vertical movement (pixel) = .095 cm]. The K_3 gain was selected to provide good control without being overly sensitive. The selected value results in an overall stick to display scaling of 1.3 cm of cursor motion per degree of potentiometer rotation (3.33 units/deg x 4 units/unit x 0.95 cm/unit), since the plant dynamics have a static gain of 1.0.

The real-time tracking loop software is free-running (i.e., the iteration rate is not directly controlled by clock interrupts). As a result, the full 21 msec time frame is used for computation of the new cursor position given the sampled stick value. Despite the fact that the tracking loop is free-running, the iteration rate (and accordingly, the frame time and trial length) varies by less than 3 percent within or across trials. A trial is flagged as invalid if the trial length varies by more than 5 percent from the minimal 180 sec value.

No external forcing function is applied to the tracking loop. The unstable dynamics are simply excited by human tracking remnant and by noise in the stick digitization process. If the subject loses control and the cursor travel reaches the edge of the display, it is automatically reset to display center and the subject continues tracking. The active area of the display is ±9.5 cm.

Calculation of the average absolute tracking error

$$E = \sum_{i}^{n} |e_i|/n$$

and the number of control losses is based on the sampled values from each time frame. The software permits the user to break the trial up into 1 sec segments for detailed analysis of tracking performance. Thus, at the finest level of resolution, the average absolute error scores are based on 47 samples of instantaneous error.

Testing Procedure

Major training effects are eliminated with six practice trials at each loading level. However, performance stability improves when practice is extended to 12 trials. Training and test trials last 3 minutes each. Instructions should specify that the cursor be kept centered in the target area as much of the time as possible, and that allowing the cursor to completely leave the screen is to be avoided. Subjects are given 10 seconds to gain control over the cursor before timing and error scoring begins for the trial. The controlled element is intended to have the appearance of a pursued aircraft, and the target area is marked by left and right notches at the screen's center. Subjects view the screen from a distance of roughly 60 cm.

CTS INTERVAL PRODUCTION TASK

Description

The CTS Interval Production task, based on a task developed by Michon (1966), is designed to draw upon human information processing resources associated with response timing. The operator is required to generate a series of equal time intervals by producing a motor response (finger tapping) at a consistent rate within the range of one to three responses per second. Variability of the duration of produced intervals is the main index of performance. Unlike the other CTS tasks, the interval production task has only one demand level.

Dependent Measures

There are two measures of tapping performance, the standard deviation of interval durations and the IPT variability score. Michon (1966) suggested the IPT variability score because it corrects for the partial dependence of error magnitude on interval duration. IPT variability is calculated by the following formula,

$$\frac{N}{T} \sum_{i=1}^{N} \left| \Delta t_i \right|$$

where N is the total number of intervals produced, T is the total time over which data are collected, and Δt is the difference between successive intervals. A lower IPT variability score indicates more temporally regular tapping and better performance. Typical variability scores range from 10 to 40.

Testing Procedure

Fifteen minutes of practice tapping are adequate for training. Subjects should be instructed to tap between one and three times per second, and to become as automatic as possible. Responses are made on a specially designed tapping key. Initially, six 30-second practice trials should be run to allow the subject to establish and maintain an acceptable tapping rate. The experimenter may need to "coach" the subject during these trials. It is best if a 2 taps/second rate is established early in training so that subsequent drift in tapping rate does not lead to unacceptable data. Four 3-minute trials should then be completed to provide sufficient practice, for a total of 15 minutes of training.

Appendix
CRITERION TASK SET HARDWARE AND SOFTWARE

The CTS (V1.0) is implemented on a commercially available microcomputer system with a minimum of additional custom-built hardware. The total cost of all necessary equipment is approximately $1500. An equipment listing is provided below.

Commodore 64 microcomputer
Commodore 1541 disk drive
Commodore C1526 printer (or MPS-801)
Monochrome experimenter's monitor (Panasonic WB5200 or equivalent with 75 ohm loop-through and female BNC video input connector)
Commodore 1702 color subject's monitor (substitute not recommended)
Experimenter's video monitor switch and cables (custom)
Four button response keypad and cable (custom)
Tapping key and cable (custom)
Rotary tracking control and cable (custom)

The equipment is configured as a subject's test station and an experimenter's station as shown in Figures 12 and 13. The experimenter's station includes the microcomputer, monochrome monitor, disk drive, printer, and video monitor switch. The switch permits the experimenter to blank the subject's screen while selecting CTS tasks and experimental conditions. The disk drive is used to load the CTS software and to store subject data.

The subject's station includes the color monitor and three special response input devices. The tapping key is used for the CTS Interval Production task. The rotary control is used for the CTS Unstable Tracking task, and the four button keypad is designed to be compatible with the spatial layout of the CTS Probability Monitoring task. The six remaining CTS tasks are binary choice response tasks which use the two keys to the extreme left and right on the four button keypad as input devices.

The software for the CTS is written primarily in BASIC to run on the Commodore 64 computer. The majority of the programs are compiled to improve

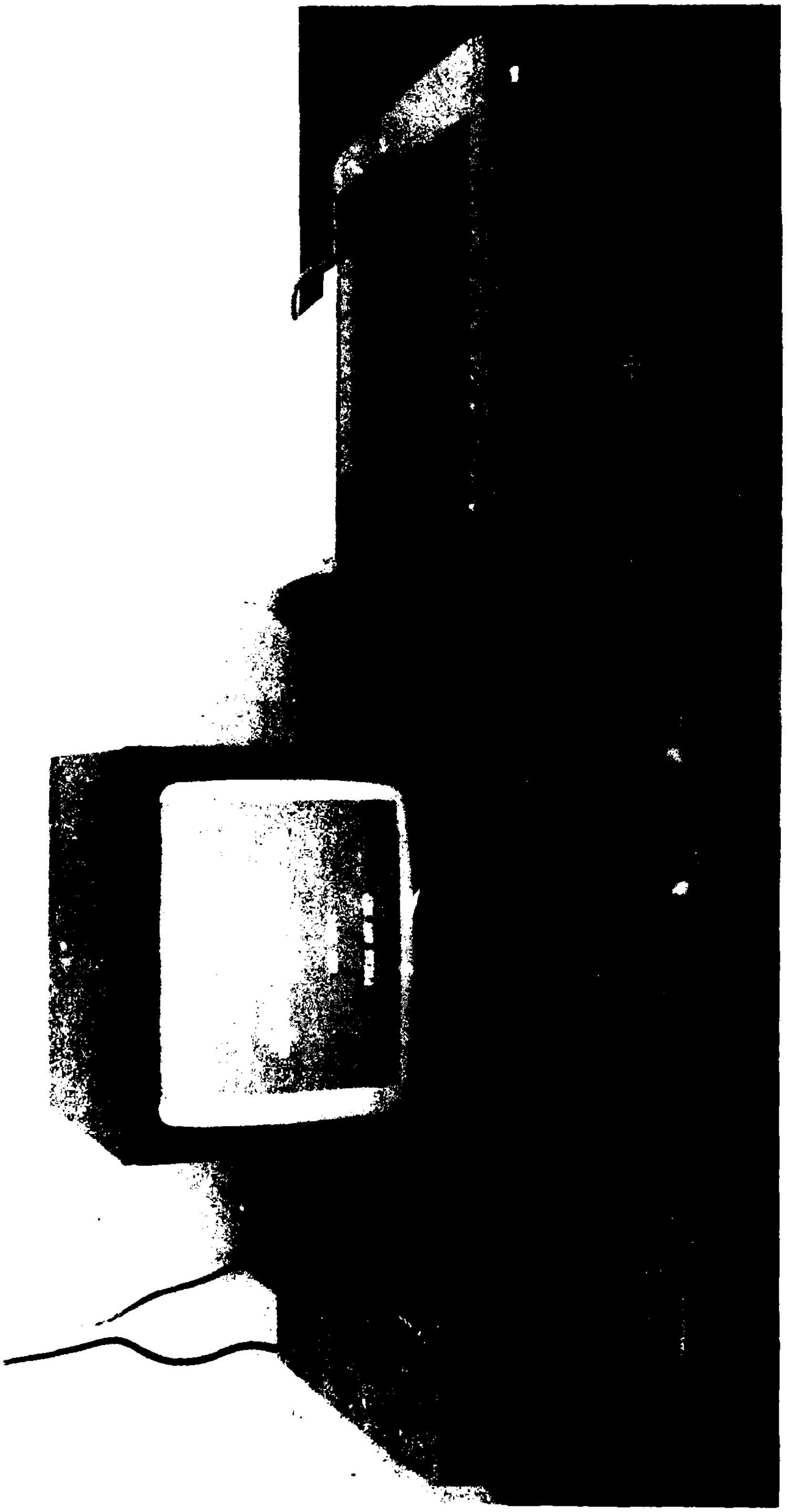


Figure 12. CTS Experimenter's Station

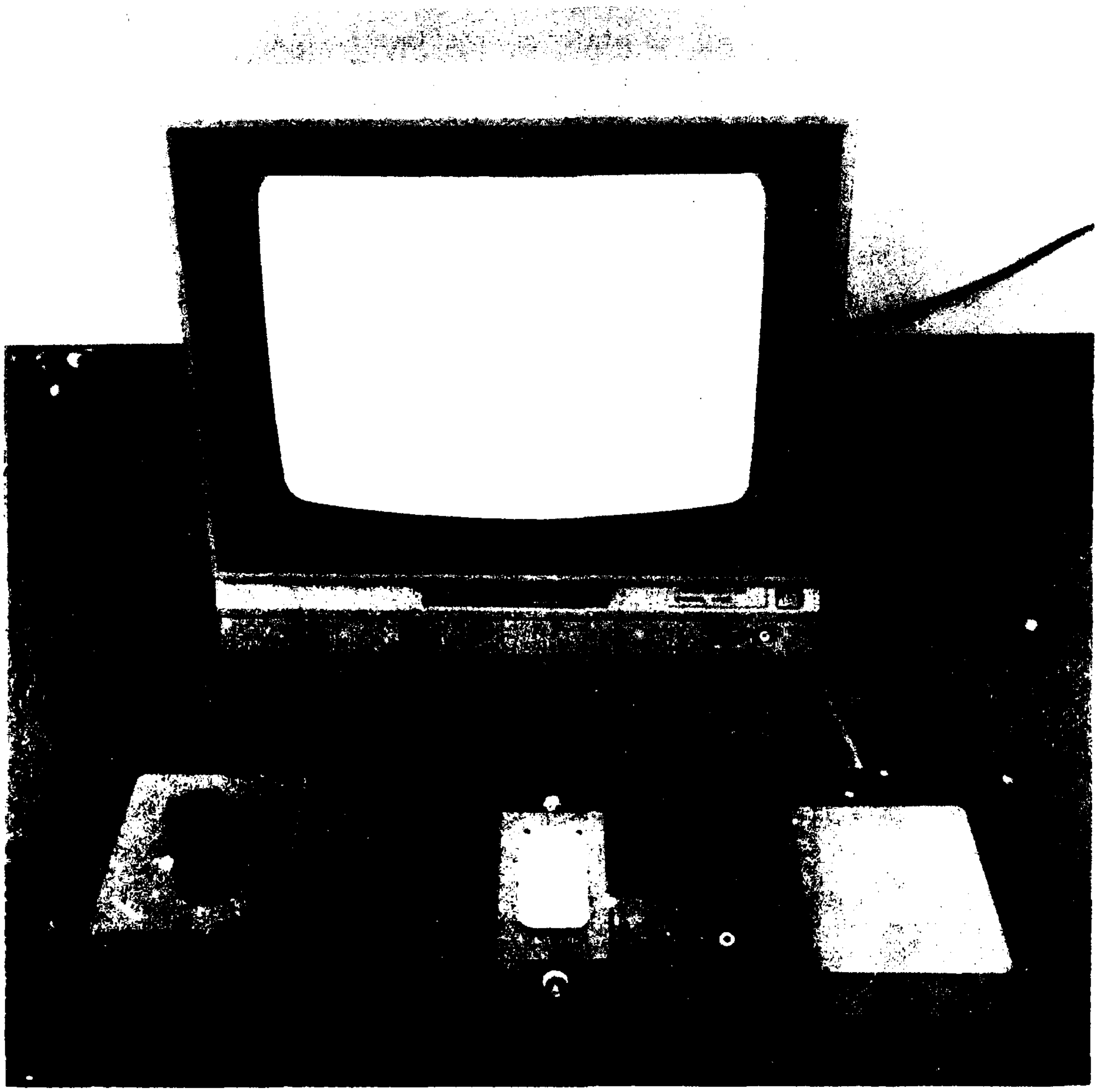


Figure 13. CTS Subject's Station

execution speed and efficiency. The reaction time measures used in seven of the CTS tasks are recorded in milliseconds with a resolution of ±1.5 milliseconds. An exception is the CTS Probability Monitoring task in which responses are timed to the nearest .10 second. The overall iteration rate of the real-time loop in the CTS Unstable Tracking task is 48 Hz.

The CTS software is structured to minimize experimenter familiarization and training requirements. Standardized, self-explanatory menus are used for all tasks to simplify trial preparation and data handling activities. Once task software is loaded into the computer, initial menus permit the experimenter to select training or test conditions and specific loading levels on the task. Options are also provided to test the response device for the task, to analyze previously stored raw data, and to display correct responses along with each stimulus presentation when required for training. Furthermore, explicit prompts are given to sequence the user through the menus and to ensure accurate insertion of subject and test condition identifiers. Following data collection, additional menus allow the experimenter to examine the new data in a "quick look" mode; calculate summary statistics; and store or print a detailed, time-based record of all stimuli presented and subject responses.

Tests of the CTS hardware and software under actual experimental data collection conditions have indicated that the combined system is highly reliable. No hardware failures were experienced during approximately 2500 3-minute test trials run over a 2-week period. In addition, experimenter error was minimized by the user-friendly software design which limited cases of irretrievable data loss to .2 percent of all test trials.

REFERENCES

Baddeley, A. D., 1968, A 3-Minute Reasoning Test Based on Grammatical Transformation, Psychonomic Science, 10(10), 341-347.

Chiles, W. D., Alluisi, E. A., and Adams, O. S., 1968, Work Schedules and Performance During Confinement, Human Factors, 10(2), 143-196.

Craik, F. I. M. and Tulving, E., 1975, Depth of Processing and the Retention of Words in Episodic Memory, Journal of Experimental Psychology: General, 104(3), 268-294.

Hunter, D. R., November 1975, Development of an Enlisted Psychomotor/Perceptual Test Battery, AFHRL-TR-75-60, Air Force Human Resources Laboratory, Personnel Research Division.

Jex, H. R., McDonnell, J. D., and Phatak, A. V., 1966, Critical Tracking Task for Manual Control Research, IEEE Transactions on Human Factors Engineering, HFE-7, 138-145.

Michon, J. A., 1966, Tapping Regularity as a Measure of Perceptual Motor Load, Ergonomics, 9(5), 401-412.

National Research Council, 1983, Research Needs for Human Factors, Committee on Human Factors, Commission on Behavioral and Social Sciences and Education, National Academy Press, Washington, D.C.

Posner, M. I. and Mitchell, R. F., 1967, Chronometric Analysis of Classification, Psychological Review, 74(5), 392-409.

Roget, P. M., 1923, Roget's International Thesaurus of English Words and Phrases, New York: Thomas Y. Crowell Company.

Shingledecker, C. A., Acton, W. H., and Crabtree, M. S., 1983, Development and Application of a Criterion Task Set for Workload Metric Evaluation, Second Aerospace Behavioral Engineering Technology Conference Proceedings, Society of Automotive Engineers, 43-49.

Shulman, A. I., 1974, Memory for Words Recently Classified, Memory and Cognition, 2(1), 47-52.

Sternberg, S., 1969, Memory Scanning: Mental Processes Revealed by Reaction Time Experiments, American Scientist, 57, 421-457.

Wickens, C. D., 1981, Processing Resources in Attention, Dual Task Performance, and Workload Assessment, Technical Report EPL-81-3, Engineering Psychology Research Laboratory, University of Illinois at Urbana-Champaign.

☆ U.S. GOVERNMENT PRINTING OFFICE 1986 - 559-006/10229

END

FILMED

6-85

DTIC